Value on the D-Line

Sarah Hu and Evan Light Rake

Problem

Which position on the defensive line has the biggest impact when on the field?

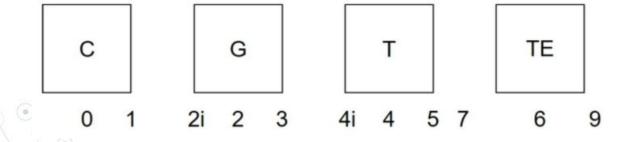
Data cleaning

☐ Create a metric for determining value of each position

Data Cleaning

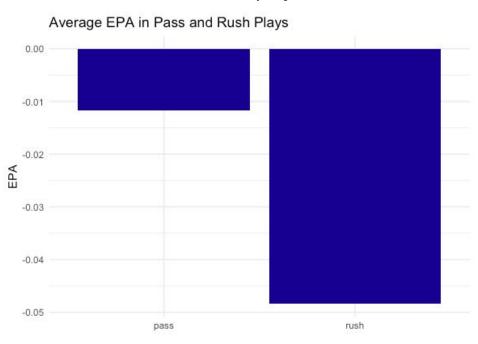
- ☐ Which positions should be considered on the D-Line?
 - Filter out wide receivers and offensive tackles

- ☐ Which players are in proximity to the ball
 - Filter out defenders that don't set up on the d-line



☐ Split data by pass plays and rush plays

- The differing offensive strategies require different defensive countermeasures; some stats are more valuable on certain plays.



	Мо	e readable stats:		
		uTimeLeft ☐ Seconds left in the game mutated from 'Quart	er' and 'TimeLeft'	
	YardsToGo Field position defined by yards to the endzone mutated from 'SideOfField' and 'StartYard' UTechniqueName Updating 'TechniqueName' with 'SideOfBall' to create a numeric representation of where defensive lineman set up relative to the center			and 'StartYard'
				ration of where
		 URUNDirection □ Updating categorical 'RunDirection' variables to a numeric representation that could be matched up with 'uTechniqueName' 		
		DistanceToRush □ Distance from lineman to rush direction in techniques: 'uTechniqueName'- 'uRunDirection'		
•	••••••••••••••••••••••••••••••••••••••	CloseToRush Binary variable for if 'DistanceToRush' is withi	n two techniques of the rush	

0

Creating a Value Metric



Why EPA

- ☐ Sums up expected value of each play
- Takes into account various factors
 - Field position
 - **→** Down
 - Distance
- Points are wins
- ☐ Negative EPA → Better Defensive Play

Regression

□ Relative value of each action (sacks, tackles, interceptions, etc) against EPA

Passing Plays

```
Call:
lm(formula = EPA ~ FumbleByPasser + SoloTackle + Pressure + SoloSack +
    AssistedSack + PassBreakup + Interception, data = pass)
Residuals:
     Min
                                         Max
-11.0463 -0.7220
                   -0.1874
                             0.8059
                                      6.7401
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)
                                      7.369 1.76e-13 ***
                 0.05402
FumbleByPasser1 -3.29229
                            0.07746 -42.505 < 2e-16 ***
SoloTackle
                 0.20277
                            0.08354
                                      2,427
                                              0.0152
Pressure
                -0.24648
                                    -8.951 < 2e-16
                -1.23129
                            0.10907 -11.289 < 2e-16 ***
SoloSack
                -1.30823
AssistedSack
                            0.14445 -9.057 < 2e-16 ***
PassBreakup
                -0.94381
```

-9.865 < 2e-16 ***

Interception

-4.42920

Rushing Plays

0.064204 -2.371

0.0178 *

```
Call:
lm(formula = EPA ~ SoloTackle + AssistedTackle + ForcedFumble +
    RecoveredFumble + CloseToRush + IsRushing + Pressure, data = runsdata)
Residuals:
    Min
            10 Median
-8.5422 -0.4156 -0.0732 0.4405 5.5516
Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)
                -0.036024
                            0.005502 -6.548 5.94e-11 ***
SoloTackle
                           0.020342 -11.087 < 2e-16 ***
                -0.225535
AssistedTackle -0.192299
                           0.019970 -9.629
                                             < 2e-16
ForcedFumble
                -2.751110
                           0.171955 -15.999
                                             < 2e-16 ***
RecoveredFumble -4.810233
                           0.192028 -25.050 < 2e-16 ***
                -0.034058
                           0.014843 -2.295
CloseToRush
                                              0.0218 *
IsRushina
                0.455901
                           0.023610 19.310
                                             < Ze-16 ***
```

-0.152207

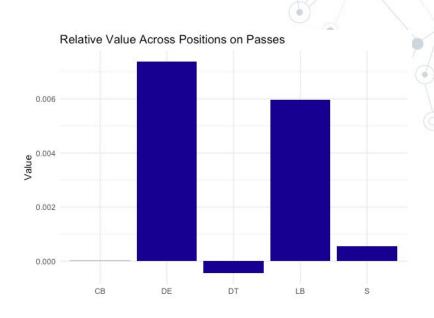
Pressure

Findings



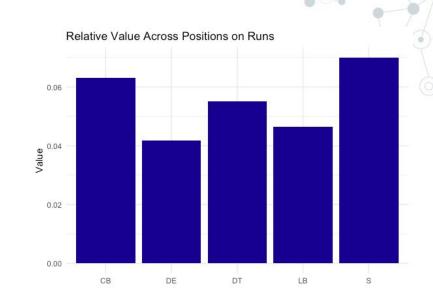
Passing Plays

- Defensive Ends and Linebackers dominate
- → Defensive Tackles have **negative** value on pass plays
- Interceptions and fumbles created make for largest impact
- Sacks and pressures have less value

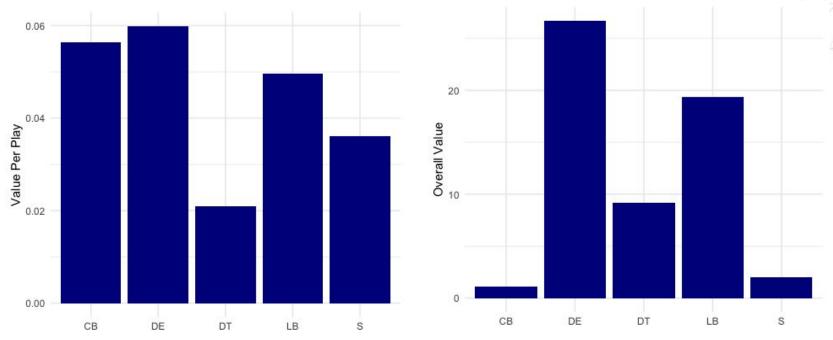


Rushing Plays

- Safeties and Cornerbacks have slightly higher values than other positions
 - Extreme emphasis placed on fumbles
- Defensive Tackles are worth more than Defensive Ends and Linebackers
- ☐ Less variance of positional value

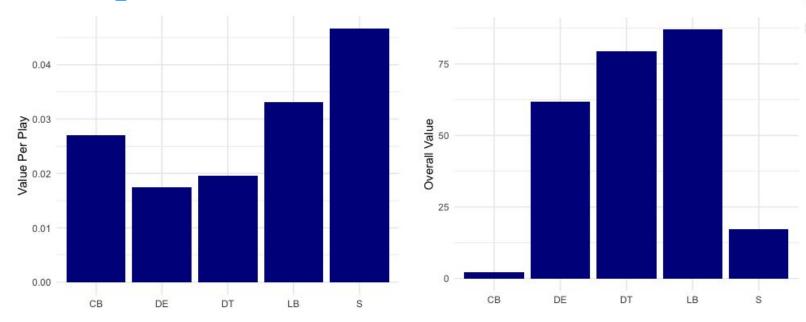


Special Situations: 4th Downs



Defensive Ends continue to stand out on 4th down plays, having both the highest value per play provided and overall value amongst all positions. Despite the large value per play provided by CBs, they are rarely present on the line and provide little overall value compared to the other positions.

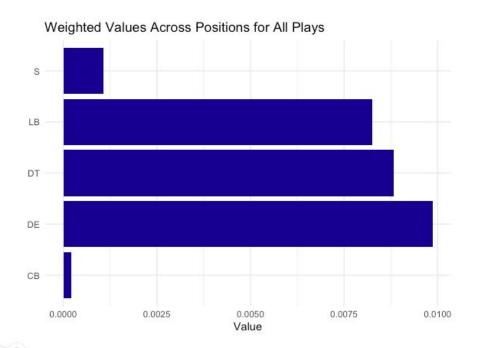
Special Situations: Red Zone



The value provided by defenders in the red zone shows that Safeties and Linebackers make a big difference near the endzone; these positions have the ability to make game changing plays and in critical spots. Although defensive ends still hold their own, linebackers really stand out in the red zone.

Combining Runs and Passes

- ☐ Weight value across positions based on frequency of pass and rush plays
- Final value metric for comparing positions



Insights

- Defensive Ends generate the most value
 - Slight edge over defensive tackles and linebackers given the way the game is shifting towards offensive pass efficiency
- Cornerbacks and Safeties create little to no value when placed on the D-Line
- ☐ Fourth downs aren't different from other plays
 - Defensive ends still provide most value
- Red Zone puts emphasis on the backfield

Reflection

Assumptions

- Safeties and Cornerbacks are playing out of their true positions
 - These positions require different skills than defensive lineman, leading them to have a disproportionate *value per play* versus *overall value* in this dataset consisting 'OnFieldPosition' of only LB and DL
- CloseToRush
 - A player is within proximity of the rusher if he is within 2 "techniques"

Future Improvements

- Seek data about matchups (ex: O-Line vs D-Line, CB vs WR)
 - Insight into skill discrepancies
 - Indicator of double team
 - Indicator of defensive/offensive formation
- ☐ Incorporate player tracking data / ball tracking data
 - Indicates more than just where the players are at the snap
 - Follow proximity of player to location of ball